7.電析のクロスオーバー、脳内シナプスの自己組織モデルについて(「パターン形成、運動及びその統計」研究会,研究会報告)

宇佐見, 義之; 寺門, 弘訓; 今野, 紀雄; 長谷, 隆

7. 電析のクロスオーバー、
脳内シナプスの自己組織モデルについて

宇佐見義之　神奈川大工、　寺門弘訓　鉾田農高

今野紀雄　室蘭工大、　長谷隆　静岡大工短

はじめに

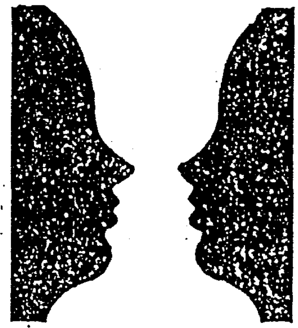

図1

パターン認識とパターン形成の問題の類似性について初めて議論したのはHaken であると思われる。例えば、図1の絵はパターン認識の不定性を示すものとしてよく知られているが、黒の領域に着目すれば花瓶のように見え、翻って白の領域に意識を向けると二人の人がキスをしているように見えてくる。つまり私達の注意の向ける具合いによってもとの対称性が破れることになる、とHaken は指摘する。[1-2]　近年彼らはより具体的に“内容番地連想記憶“モデル (content-addressable associative memory) の一種としてテストベクトルが元の教師ベクトルに q^2, q^4 のポテンシャル力で近ずく模型を提案して、このような認識に関する対称性の破れを表すことを試みている。

我々は今回、各々のシナプスの結合荷重が二重安定点、更には多重安定点のポテンシャルを持つ場合にそれらが結合された時の全体のネットワークがどのような特性を示すかについて考察を行った。　以下に示すように、この場合には Haken らが主張するのとはまた別の意味で“パターン認識“の問題が物理で研究されている“パターン形成“の問題と類似性を帯びる。

二重安定点をもつシナプスが作るニューロン群のネットワーク特性

ニューラルネットワークのカテゴリーとしてはマッカロ・ピッツ型の興奮特性を持つニューロンに Hebb ±型の学習則をもつ教師なし自己学習型ニューラルネットワークを考える。　マッカロ・ピッツ型のニューロンとは x を入力とした時その出力 f(x) が階段関数である場合を指す。また Hebb 型の学習則とは今の場合、ニューロンが興奮して発火した時、そこのシナプス荷重（信号の伝達効率）が増強する、というものをさす。　更に Hebb ± 型の学習規則とは興奮したニューロンのシナプス伝達効率は増強するが、逆に入力の無いニューロン間のシナプス伝達効率は下がるというものである。[3]　今のモデルは対象

としては、脳のおもに視覚系のパターン認識（短期記憶）、及び視覚的な情報として保存されているであろう長期記憶、を想定している。 しかしこの研究は現実の脳の機構をモデル化するという試みではなく、むしろこのようなモデルの機能を考察することによってニューラルネットのもつパターン認識とその長期保持などの機構についての洞察を得ることを目的としている。

Hebb± 型の学習方程式は通常

$$(d/dt)\ W_{ij} = -\alpha W_{ij} + \eta (X_i - 0.5)\, f\left(\sum_i W_{ij} X_i - \theta\right) \tag{1}$$

Xj	1 or 0
Wij	シナプス結合荷重
α	減衰係数
$f(\theta)$	階段関数

と表され、右辺の第２項が学習の為の入力の部分を表現している。[4] この学習方程式の第１項は、W_{ij} の potential を $V_{ij}(W_{ij})$ として

$$(d/dt)\ W_{ij} = -(dVij/dWij) + \eta (X_i - 0.5)\, f\left(\sum_i W_{ij} X_i - \theta\right) \tag{2a}$$

$$V_{ij}(W_{ij}) = (1/2)\alpha\ W_{ij}^2 \tag{2b}$$

ととったことに相当している。 この項は入力 X_i が無い時に $W_{ij} \rightarrow 0$ となるような忘却の効果を表すものである。 以後、一つのシナプス結合 W_{ij} の potential に的を絞って議論し、その為添え字の i, j を省略する。 筆者らの提案は、シナプス結合 W のポテンシャルを次の double miminum を持つものを考えると、入力が無くなっても覚えたことを保持することになるのではないかということである。

シナプス結合が二重安定点を持つ場合

Potential を次のようにW^4型にとると

$$V(W) = -aW^2 + bW^4 \tag{3}$$

学習の時間変化は図２に示されるようになることが期待される。

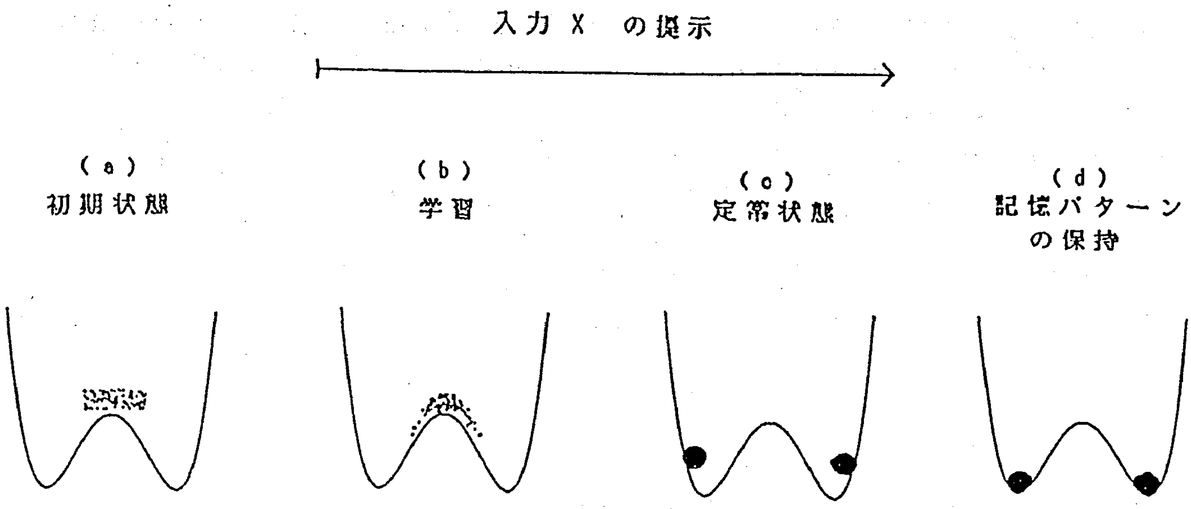

図2　2重安定点をもつポテンシャル上での学習過程。　ランダムに散らばっているシナプス結合荷重がパターンの要素(右)とそれ以外の雑音(左) に分かれる。

このように (2)の学習方程式の potential の項が複雑になった場合はパターンの入力と Potential Dynamics の競合が起こる。　例えば potential の項が $(d/dt)W = \gamma W$ で不安定化する場合 ($t \sim 0$)、t=0 で $|W| < W_0 < 1$ の初期条件 に対して $W(t) = W_0 \exp(\gamma t)$ なので potential による力に対抗してパターンを分離させる為には $W_0 (\exp(\gamma t) - 1) < \eta$ (あるいは $\gamma t < \ln(\eta / W_0 + 1)$) の関係を満たした η (時刻 t 後の信号の強さ)が必要である。　詳細は省くことにして、今 (2) 式の学習方程式のポテンシャルが (3) の様な二重安定点を持つ場合を考え、その学習過程を示そう。
例えば、potential 項の不安定化の定数でスケールした学習方程式は

$$(d/dt)\, W = W - \gamma_1 W^3 + \eta (X_i - 0.5)\, f\, (\Sigma W_{ij} X_j - \theta)\, \delta(\tau - n\tau_0) \quad (5)$$

となる。 初期条件、各パラメーターは、t=0, での W のばらつきを $|W_0| < 0.3$, 入力の時間間隔 τ_0 を不安定化の特徴的な時間の 1/10 、とし先の評価 $\gamma t < \ln(\eta / W_0 + 1)$ を $0.10 < 0.51$ で満たしている。学習は A, B というパターンを τ_0 の時間間隔で交互に入力した (n=0, 1, 2、・・・・)。

図3に学習するシナプス結合荷重の時間変化、図4にWの軌跡とポテンシャルの関係を示した。　図4の方から見ると、 t=0 において $|W_0| < 0.3$ でばらついていたシナプス結合荷重 Wが入力の要素である A (B) という部分だけポテンシャルの右側に押しやられ、A (B) の要素以外のWは雑音として dou-

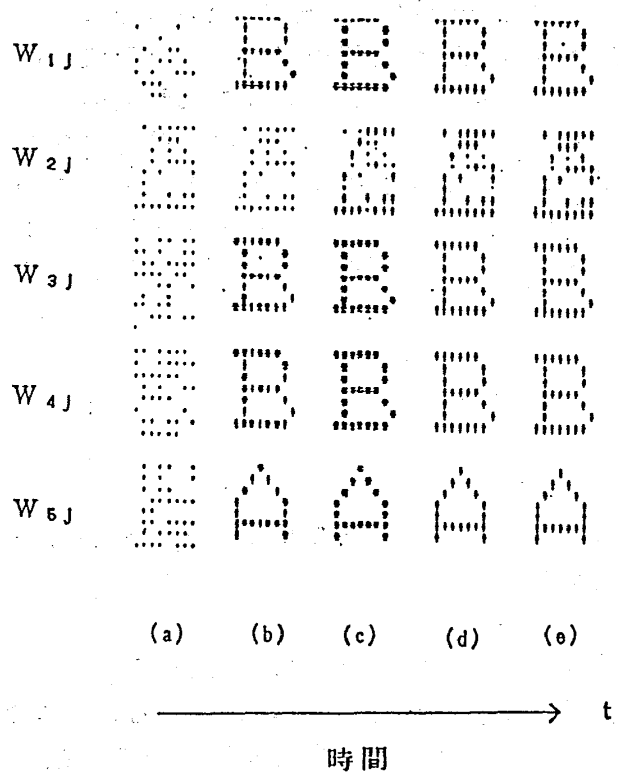

図3　結合荷重 W_{iJ} のJ番目の値の時間変化。 A, B というパターンを分離して記憶している。

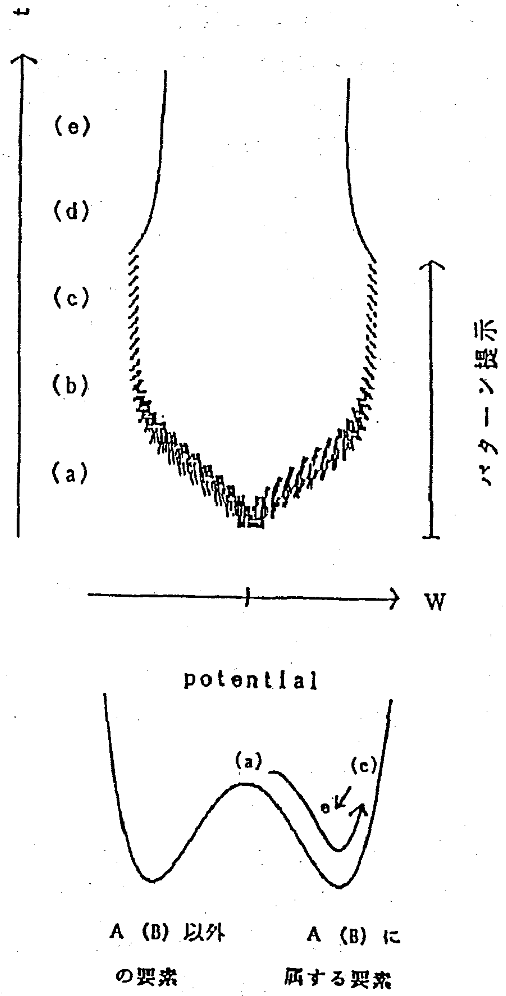

図4　W_{iJ} のトラジェクトリー

図3に学習するシナプス結合荷重の時間変化、図4にWの軌跡とポテンシャルの関係を示した。 図4の方から見ると、 t=0 において $|W_0| < 0.3$ でばらついていたシナプス結合荷重 Wが入力の要素である A (B) という部分だけポテンシャルの右側に押しやられ、A (B) の要素以外のWは雑音として double minimum potential の左側に移行する。 各 W の軌跡は時間が経つに従って、 (a)→(e) と変化するが (b), (c)は入力とポテンシャルの吸引力がつりあった定常状態、 (d), (e) は入力が途絶えたあとのパターンの保持をにWがあることを示している。 また図3を見ると確かに t=0 でランダムにバラついている結合荷重Wが学習により A, B というパターンに分離し、かつ入力が途絶えた後もそのパターンを保持していることがわかる。

以上、 Hebb 則の学習方程式の potential 項が2重安定点を持つ場合の Neural Netork について簡単に紹介した。 紙面を配慮して触れなかったが、このような neural network のパターン分離（認識）が物理の“パターン形成“の問題とと非常に似ていることはおわかりいただけると思う。 q^2, q^4 型の potential 上での秩序形成は鈴木によって詳しく調べられているので参考になるが[5]、今の場合重要なのはむしろ初期での分離が重要のようである。またこのモデルで興味深い点は、 講演でも触れたように potential をもっと複雑にすると、様々な記憶のプロセスが記述できるようになる可能性にあるが、詳しくは別の機会に述べる予定である。

おわりに

本来の脳研究の醍醐味は生命体の認識機構、認識の情報科学、人工知能の問題などとの接点にあり、この場合の戦略が
(1) 神経生理学や認知心理学のデータを参考にしつつ
(2) 外界の無限性と生体の有限性の関係について具体的な定式化を行う
ことにあると言える[6]、とするなら本研究は(1)の一部分を埋めているに過ぎない。 いわゆる人工知能研究がこれらから遠いところに目を向けている現在、筆者は(2)の立場からの考察も進めているが、それはまた別な機会に論じる。ただ、科学を含む様々な事柄が脳と外界の接点を調節しようとする脳自身の発展から説明できると説く養老の視点はまさしく無限の多様性をもつ世界と有限である生命体の関係のダイナミズムを考える立場にあり、おおいに参考になろう。[7] 要は脳の構造に注意しながら、その機能を論じることに将来的な発展があろうということである。 また、迂遠なことのように思われるかもしれないが、生命、生命体の認識などを論じる場合、郡司、橋爪、ハート、ウィトゲンシュタイン らが展開した“言語ゲーム“的な見方[8,9]、つまりモデルを記述する際の我々の言語について問い直さなければならないということがこの場合重要な役割を果たすように筆者には思えるがこれについても後日詳しく論じる予定である。 しかしこの際にも具体的に計算可能なモデルの提出が重要なのはいうまでもない。

本講演の電析に関する部分は長谷との共同研究で、これについては論文(10)を参照のことを付記しておく。

文献

1) Pattern formation and pattern recognition, H.Haken (ed) Pattern formation by dynamical systems and pattern recogition. Springer Berlin Deidelberg New York (1979).
2) Oscillations in the Perception of Ambiguous Patterns, T.Ditzinger and H.Haken, Bio.Cybern.61,279(1989).

3) この考えは筆者らが独立に考案したものだが、時期的には仁木らによって先に発表されている。 誤差逆伝播学習アルゴリズムはヘッブの学習アルゴリズムと共存できるか？, 仁木和久, 山田誠之, 信学技報, NC89, 49(1990).

4) 入力の時間特性についてはここでは詳しく述べない。

5) M. Suzuki, Adv. Chem. Phys., 46, 195(1981).

6) 認識の情報科学への計算論的アプローチ、安西祐一郎、人口知能学会誌、3, 248(1988).

7) 唯脳論、 養老孟、青土社、(1989).

8) 言語ゲームと社会理論、橋爪大三郎、頸草書房、(1985).

9) Time reverse automata pattern generated by Spencer-Brown's modulator: Invertibility based on the autopoiesis, Y. Gunji and T. Nakamura, to appear in Bio. System.

10) Diffusion-Limited Aggregation in Coupled Diffusion Fields, Y. Usami and T. Nagatani, J. Phys. Soc. Jpn., 59. 474(1990).